COMPUTER AIDS FOR AUTHORING TESTS

J. Peter Kincaid and Richard Braby
Training Analysis and Evaluation Group
Orlando, Florida

and

Wallace H. Wulfeck, II
Navy Personnel Research and Development Center
San Diego, California

ABSTRACT

Computer routines developed to help in authoring and editing textual training materials were modified to aid authors of tests. The Navy's Computer Readability Editing System (CRES) aids in producing comprehensible text by flagging uncommon words and awkward sentences, suggesting replacements for awkward words or phrases, and giving the readability grade level. Additional routines were developed, based on the Instructional Quality Inventory, specifically for multiple choice and true/false test questions. These new routines calculate readability grade level of test questions, and flag some kinds of awkward or incorrect test item construction. The CRES routines, including the new test item features, are intended to be used as part of a computer-based publishing system. Our initial effort to provide feedback to authors of tests has convinced us that the general approach is viable and many new useful features could be added.

Computer routines to aid authors in developing tests are now feasible. Some routines have already been developed. For example, Roid and Finn (1978) have demonstrated routines for generating multiple choice test items from text passages. Additional routines to aid authors in test items would complement the Roid and Finn routines, or could be used separately.

COMPUTER READABILITY EDITING SYSTEM (CRES)

This paper describes a new version of the Computer Readability Editing System (CRES) modified to provide aid in authoring tests. Like the original version of CRES, it is designed to make written material easier to understand. The original version of CRES, developed specifically for text, is documented in Kincaid, Aagard, and O'Hara (1980); Kincaid, Cottrell, Aagard, and Riseley (1981); Kincaid, Aagard, O'Hara, and Cottrell (1981); and Braby and Kincaid (1981-82).

The basic configuration of the original CRES is shown in the flow chart contained in figure 1. The steps depicted for using the CRES as part of a computer-based publishing system are:

- Choose program options.
- Author or typist enters text (typically 500 to 2,000 words).
- Text is analyzed and printed out and shown on the display.
- Author revises text prompted by computer-generated suggestions (see figure 2).
- Revisions are entered and stored.
- Revised text is again analyzed by the computer to obtain readability grade level and check for keying errors.
- Editor approves text which is then stored for final camera-ready printout (for example, using a daisy wheel printer).

The basic features of the original system include those which:

- Flag uncommon words - those not on a carefully prepared list of 4,300 common words or a series of supplementary technical word lists, each about 100-200 words.
- Flag long or awkward sentences - those with passive voice or double negatives.
- Suggest replacements for awkward words and phrases.
- Provide the readability grade level according to the Department of Defense standard, the Flesch-Kincaid formula.

In addition, the system flags misspelled words.

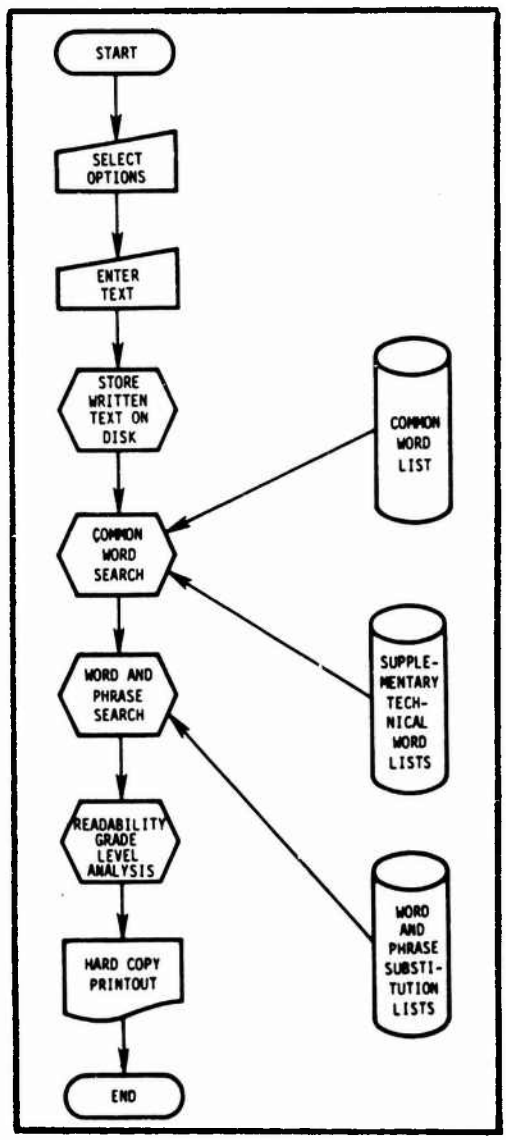


Figure 1. Flowchart Showing Phases of Operation of the Computer Readability Editing System (CRES)

INSTRUCTIONAL QUALITY INVENTORY (IQI)

The Instructional Quality Inventory (IQI) is a set of procedures for quality control of instructional development, designed to parallel and supplement the Instructional System Development (ISD) process. The IQI procedures can also be used to evaluate existing instruction, and can be used as evaluation or acceptance tools for instructional programs obtained through contract. A series of Navy reports document the IQI (e.g., Ellis, Wulfeck, Merrill, Richards, Schmidt, and Wood, 1978).

Ellis, Wulfeck, and Fredericks (1979) describe a series of steps in using the IQI:

- Classification. The IQI procedures are based on a system for classifying objectives, test items, and instructional components. Classification is determined according to: (a) what the student is required to do with the information to be learned, and (b) what type of information the student is learning.

- Assuring Objective Adequacy. Since good instruction depends on careful specification of learning objectives, the first IQI procedure is to assure the adequacy of objectives. This is done by classifying each objective, and judging whether or not it accurately reflects the intended student performance after training.

- Constructive, Consistent, and Accurate Tests. The next IQI step is to make sure that tests accurately measure progress toward the objectives. This is done by assessing the consistency between each test item and its associated objective, and the adequacy of the item. Essentially, each test item must be classified in the same way as its objective and must be adequately constructed.

- Keeping Presentations Consistent. Instructional presentations contain various components, including statements of material to be learned, examples, and practice. For consistency, different combinations of presentation components are required depending on the classification of the objective.

- Applying Adequate Learning Principles. The final IQI step is to make sure that each required presentation component is adequate according to psychological principles of learning.

The IQI is, among other things, a checklist for preparing quality tests.

This paper describes the incorporation of a number of these items into the CRES, those which could easily be programmed.

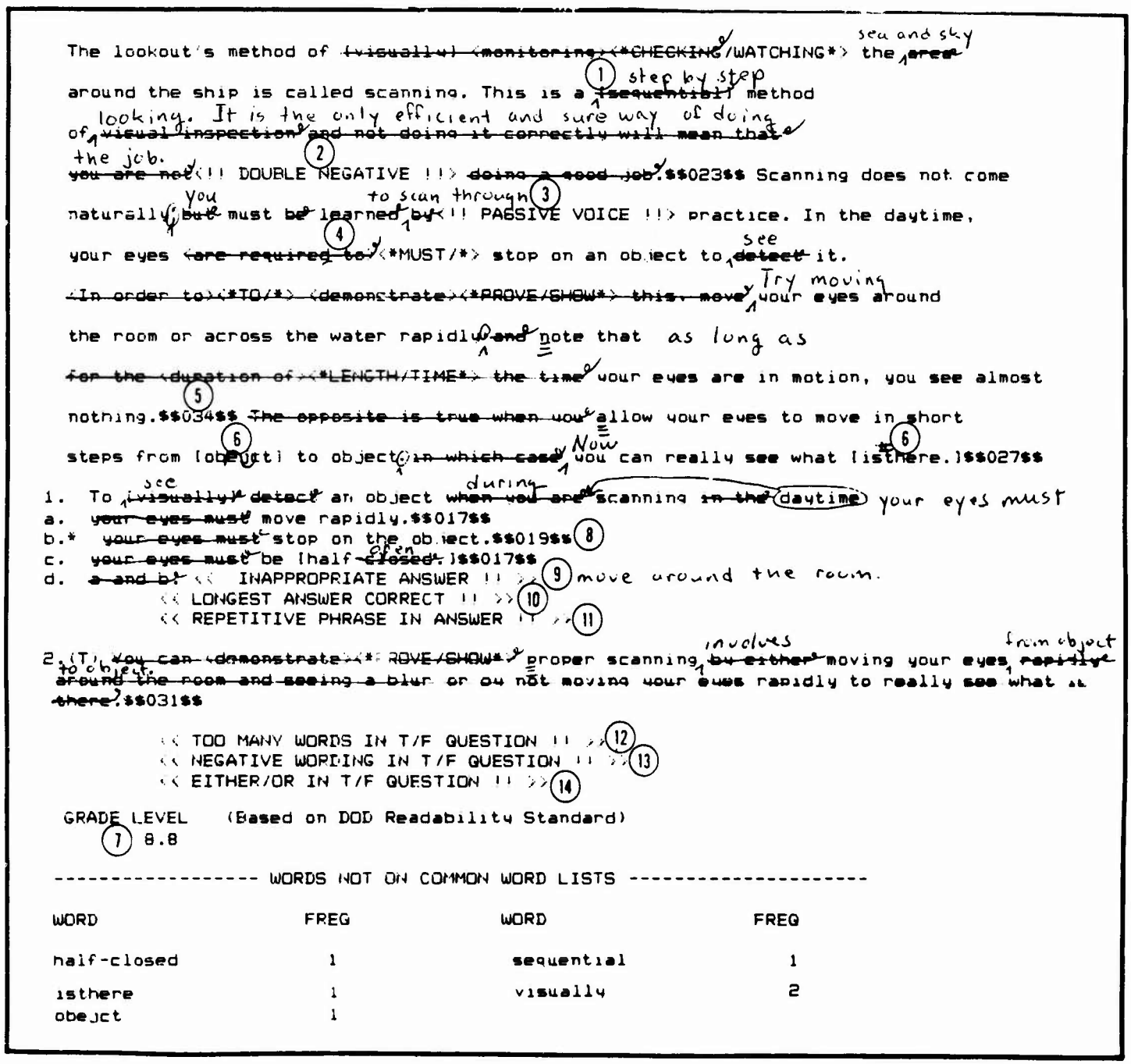
EXPANDED CRES FOR TEST ITEM ANALYSIS

Figure 2 shows a CRES analysis and illustrates both the original features of the CRES (for analyzing text) and new features for analyzing test questions. It also shows handwritten editing changes suggested by the computer analysis. The passage and questions are intended for a Navy remedial reading workbook for sailors with no higher than sixth grade reading ability.

Features of the CRES designed for text are indicated by callouts 1-7.

1. Uncommon words are flagged: in this case "sequential."

2. Double negatives are flagged: in this case "not...not."



The lookout's method of ~~[visually] <monitoring:<*CHECKING~~/WATCHING*> the ~~area~~ sea and sky around the ship is called scanning. This is a ~~[sequential]~~ (1) step by step method of ~~visual inspection and not doing it correctly will mean that~~ looking. It is the only efficient and sure way of doing the job. ~~you are not~~ (2) <!! DOUBLE NEGATIVE !!> ~~doing a good job.~~$$023$$ Scanning does not come naturally ~~but~~ you must ~~be~~ learned ~~by~~ (3) <!! PASSIVE VOICE !!> practice. In the daytime, your eyes ~~are required to~~ (4) <*MUST/*> stop on an object to ~~detect~~ see it. ~~In order to><*TO/*> <demonstrate><*PROVE/SHOW*> this, move~~ Try moving your eyes around the room or across the water rapidly ~~and~~ = note that as long as ~~for the <duration of><*LENGTH/TIME*> the time~~ your eyes are in motion, you see almost nothing.$$034$$ ~~The opposite is true when you~~ allow your eyes to move in short steps from [object] to object ~~in which case~~ = Now you can really see what [isthere.]$$027$$

1. To ~~[visually]~~ see ~~detect~~ an object ~~when you are~~ during scanning ~~in the~~ daytime your eyes must
a. ~~your eyes must~~ move rapidly.$$017$$
b.* ~~your eyes must~~ stop on the object.$$019$$ (8)
c. ~~your eyes must~~ be [half ~~closed~~ often].$$017$$
d. ~~a and b~~ << INAPPROPRIATE ANSWER !! >> (9) move around the room.
<< LONGEST ANSWER CORRECT !! >> (10)
<< REPETITIVE PHRASE IN ANSWER !! >> (11)

2. (T) ~~You can <demonstrate><*PROVE/SHOW*>~~ = proper scanning ~~by either~~ involves moving your eyes ~~rapidly~~ from object to object. ~~around the room and seeing a blur or by not moving your eyes rapidly to really see what is there.~~$$031$$

<< TOO MANY WORDS IN T/F QUESTION !! >> (12)
<< NEGATIVE WORDING IN T/F QUESTION !! >> (13)
<< EITHER/OR IN T/F QUESTION !! >> (14)

GRADE LEVEL (Based on DOD Readability Standard)
(7) 8.8

---------- WORDS NOT ON COMMON WORD LISTS ----------

WORD	FREQ	WORD	FREQ
half-closed	1	sequential	1
isthere	1	visually	2
object	1		

Figure 2. Original Test Showing CRES Analysis (With Features Noted by Callouts, and Editor's Handwritten Changes)

3. Passive verbs are flagged: in this case "be learned." A passive verb is composed of a form of the auxiliary verb "to be" plus a past participle. Language experts agree that the active verb (in this case "learn") is generally easier to understand.

4. Replacements for awkward words and phrases are suggested: in this case "must" is suggested as a replacemnt for "are required to."

5. Long sentences are flagged and the number of words in the sentence is shown between dollar signs: in this case "$$34$$."

6. Keying errors and misspelled words are flagged: in this case "obejct" and "isthere." These are listed as WORDS NOT ON COMMON WORD LISTS.

7. The readability grade level calculated according to the Department of Defense readability standard, the Flesch-Kincaid formula, is shown: in this case "8.8." This grade level is for both the passage and the two test questions. The original version of CRES gives grade level only for text. Readability grade level for multiple choice questions is calculated by using the question stem as the beginning of the sentence for each alternative.

New features of the system, suggested by the IQI, specifically analyze test questions. They are indicated by callouts 8-14.

8. Long sentences in multiple choice items are flagged. The words in both the question stem and answer are counted as a single sentence and the number of words shown between dollar signs: in this case "$$19$$."

9. Inappropriate answers to a multiple choice question are flagged: in this case "a and b."

10. If the longest answer to a multiple choice question is the correct answer it is flagged. Test-wise students use this as a clue.

11. Repetitive words or phrases in answers to multiple choice items are flagged: in this case "your eyes must." If the same word or phrase starts each alternative in a multiple choice test item, it should be moved to the stem.

12. Long true/false test questions are flagged.

13. Negative wording in true/false test questions is flagged: in this case "not."

14. Certain complex true/false test questions are flagged: those containing "either...or," "neither...nor," and "or": in this case "either...or."

Figure 3 shows the revised test. It is improved in many ways compared with the original version. Not only is readability grade level reduced (from grade level 8.8 to 4.5), but the text is easier to read and several errors have been removed from the test questions.

The lookout's method of watching the sea and sky around the ship is called scanning. This is a step by step method of looking. It is the only efficient and sure way of doing the job. Scanning does not come naturally; you must learn to scan through practice. In the daytime your eyes must stop on an object to see it. Try moving your eyes around the room or across the water rapidly. Note that as long as your eyes are in motion, you see almost nothing. Allow your eyes to move in short steps from object to object. Now you can really see what is there.

1. To see an object during daytime scanning your eyes must
a. move rapidly.
b.* stop on the object.
c. be half open.
d. move around the room.

2.(T) Proper scanning involves moving your eyes from object to object.

Figure 3. Revised Test (Grade Level is 4.5)

CONCLUSIONS

The extension of the CRES routines to aid in the development of test items appears useful. It should be noted, however, that the present effort was simply a demonstration. Several dozen automatic checks of IQI items could be added using current equipment and without a major difference in the type of computer algorithms already employed.

One entirely different kind of computer check could significantly increase the scope of computer-assisted authoring. Whereas the features described in this paper are automatic, a computerized IQI checklist could be added to the system as suggested by Spannaus (1980). For example, if the objective of a lesson is to learn nomenclature, the computer could ask the author, "Is a memory aid appropriate?" This kind of query is easier to program than the automatic checks described in this paper.

The routines described in this paper merely illustrate the value of the use of computers in the development of instructional material. We can expect many more such developments in the near future.

REFERENCES

Braby, R. and Kincaid, J. P. "Computer Aided Authoring and Editing," Journal of Educational Technology Systems, Vol. 10 (2), 1981-82. (In press)

Ellis, J. A., Wulfeck, W. H. II, and Fredericks, P. S. The Quality Instructional Inventory II. User's Manual. NPRDC Special Report 79-24, 1979. Navy Personnel Research and Development Center, San Diego, CA 92152.

Ellis, J. A., Wulfeck, W. H. II, Merrill, M. D., Richards, R. E., Schmidt, R. V., and Wood, N. D. Interim Training Manual for the Instructional Quality Inventory. NPRDC Technical Note 78-5, 1978. Navy Personnel Research and Development Center, San Diego, CA 92512.

Fredericks, P. S. The Instructional Quality Inventory III. Training Workbook. NPRDC Special Report 80-25, 1980. Navy Personnel Research and Development Center, San Diego, CA 92152.

Kincaid, J. P., Aagard, J. A., and O'Hara, J. W. Development and Test of a Computer Readability Editing System (CRES). TAEG Report No. 83, 1980. Training Analysis and Evaluation Group, Orlando, FL 32813.

Kincaid, J. P., Cottrell, L. K., Aagard, J. A., and Risley, P. Implementing the Computer Readability Editing System (CRES). TAEG Report No. 98, 1981. Training Analysis and Evaluation Group, Orlando, FL 32813.

Kincaid, J. P., Aagard, J. A., O'Hara, J. W., and Cottrell, L. K. "Computer Readability Editing System," IEEE Transactions on Professional Communications, Vol. 24, pp. 38-41, March 1981.

Roid, G. and Finn, P. Algorithms for Developing Test Questions From Sentences in Instructional Materials. NPRDC Technical Report 78-23, 1978. Navy Personnel Research and Development Center, San Diego, CA 92152.

Spannaus, T. W. "Speculations on Computer Assisted Design of Instruction." Paper presented at the National Conference on Computer Based Education, Minneapolis, MN, October 1980. (ERIC #ED200215)